DOCUMENT RESUME

ED 041 617 PS 003 033

AUTHOR	Miller, Louise B.; And Others
TITLE	Experimental Variation of Head Start Curricula: A Comparison of Current Approaches. (November 1, 1969-January 31, 1970).
INSTITUTION	Louisville Univ., Ky.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
REPORT NO	PR-5
PUB DATE	31 Jan 70
NOTE	42p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$2.20
DESCRIPTORS	Behavior Patterns, *Comparative Analysis, *Curriculum Design, *Preschool Children, *Program Evaluation, Student Development, Teacher Behavior, Teaching Techniques
IDENTIFIERS	Bereiter Engelmann, DARCEE, Demonstration And Research Center Early Education, Montessori, Project Head Start

ABSTRACT

This paper reports results of the first year of a 2-year comparative study of four curricula used for disadvantaged preschool children: Bereiter-Engelmann, DARCEE, Montessori, and Traditional (the official Head Start program). Details of the study design and procedures are contained in the abbreviated Annual Progress Report for 1968-1969 (PS 003 034). Treatment (program) dimensions were assessed by in-class monitoring of teachers and children using a time-sampling procedure, and by video-tape monitoring of teachers in their classrooms. Significant differences were found among the four curricula on a number of dimensions of behavior for both teachers and children, most of these differences being in predicted directions. Treatment effects were assessed by use of a variety of cognitive, social, motivational, perceptual, and achievement measures. Programs had significantly different effects on the children with respect to a number of variables measured, such as curiosity, initiative, arithmetic, and verbal participation. Preliminary regression analyses on the relationship between teaching techniques monitored in class and dependent variables have produced multiple R's between .229 and .419 and partial R's between - .293 and .307. No interpretation has been made, pending the inclusion of variables from the video-tape monitoring. (Author/NH)

"EXPERIMENTAL VARIATION OF HEAD START CURRICULA: A COMPARISON OF CURRENT APPROACHES"

Research Grant #CG 8199

from

Office of Economic Opportunity

PROGRESS REPORT No. 5.

November 1, 1969 - January 31, 1970

Louise B. Miller, Ph.D., Project Director

Jean L. Dyer, Ph.D., Research Associate

James M. Driscoll, Ph.D., Research Associate

Robert C. Sancomb, M.A., Graduate Research Assistant

Hanna I. White, M.Ed., Graduate Research Assistant

Stephen A. White, B.A., Graduate Research Assistant

Mary Frances Weedman, Administrative Assistant

Kay L. Proctor, Secretary

PSYCHOLOGY DEPARTMENT

UNIVERSITY OF LOUISVILLE

LOUISVILLE, KY.

ABSTRACT OF RESEARCH COMPLETED

A grant from the Office of Economic Opportunity in 1968 made possible a 2-year study in which four quite different preschool programs designed for disadvantaged children were compared.

These four programs were: Bereiter-Engelmann which emphasizes acquisition of linguistic and numerical skills by use of verbal instruction, imitation, and reinforcement, and de-emphasizes sensorial stimulation and manipulation; DARCEE which emphasizes, in addition to verbal and conceptual skills, the acquisition of attitudes and motives related to learning, using verbalization, reinforcement, manipulation of materials, and imitation; Montessori, which emphasizes development of persistence, independence, and self-discipline, in addition to conceptual skills, using sensorial stimulation, manipulation of materials, and self-selection, and de-emphasizes reinforcement and verbalization; Traditional (official Head Start Program), which emphasizes development in social and emotional areas, language skills and curiosity, using manipulation of materials, sensorial stimulation, role-playing, and self-selection, and de-emphasizes verbal instruction and reinforcement.

The study was designed to provide appropriate controls for teacher and population variables, and incorporated two control groups--a non-preschool group similar to the experimental sample and a middle-class group in a private preschool.[1]

Fourteen classes were conducted during the 1968-69 school year--two Montessori classes and four classes in each of the other program styles. Four-year-olds, randomly assigned within areas to Head Start classes, were tested in the fall after about 8 weeks of school and again in the spring at the end of the school year. Nine instruments designed to assess gains in cognitive, motivational, social and perceptual development were used. Five additional tests were administered at the end of the year primarily to assess specific skill-learning. Classes were monitored five times and also video-taped during the year to assess treatment dimensions for both children and teachers.

Results are available on treatment dimensions and treatment effects. Analyses of the relationship between treatment dimensions and treatment effects are in process.

Detailed descriptions of the design and procedures are contained in the abbreviated Annual Progress Report for 1968-69.

---

[1] The middle-class control group was not obtained until the second year; therefore, results are not yet available.

SUMMARY OF RESULTS

The results obtained on program variations during the first year of the study fall into three categories: (1) Differences among programs with respect to dimensions of classroom activity ("Treatment Dimensions"), (2) Effects of programs on children ("Treatment Effects"), and (3) Relationships between Treatment Dimensions and Treatment Effects. Teacher effects, independent of programs, also occurred but have not been assessed.

1. Programs were found to differ significantly in a number of dimensions with respect to behavior of both teachers and children, most of these differences being in predicted directions.

In the Bereiter-Engelmann classes, teachers were significantly high in verbal instruction and exemplification; DARCEE teachers were significantly high in verbal instruction; Traditional teachers were significantly high in manipulation. Bereiter-Engelmann and DARCEE children were significantly high in verbal recitation; Montessori children were significantly high in manipulation; Traditional children were significantly high in role-playing. It appears almost certain that programs will also differ significantly with respect to knowledge-of-results (KOR) for correct responses, KOR for incorrect responses, contingent positive reinforcement, requests for academic verbal performance and requests for imitation.

2. Programs had significantly different effects on children with respect to a number of the variables measured.

Bereiter-Engelmann and, to some extent, DARCEE appear to have significantly affected cognitive functioning as measured by the Stanford-Binet and the Preschool Inventory; DARCEE appears to have had considerable impact on children's motivation to achieve, persistence, resistance to distraction, initiative, and curiosity; Montessori and Bereiter-Engelmann children were significantly higher than controls on curiosity. Bereiter-Engelmann produced significantly high achievement on sentence production and arithmetic;

DARCEE was significantly high on arithmetic. According to teachers' ratings after six months, DARCEE children were significantly high in verbal-social-participation and less timid than children in other programs. Bereiter-Engelmann children were significantly less aggressive.

3. Preliminary regression analyses on the relationship between teaching techniques monitored in class and dependent variables have produced multiple R's between .229 and .419, and partial R's between -.293 and .307.

No interpretation has been made of these analyses, pending the inclusion in further regression analyses of the variables from the video-tape monitoring, such as reinforcement.

"EXPERIMENTAL VARIATION OF HEAD START CURRICULA:
A COMPARISON OF CURRENT APPROACHES"

Research Grant #CG 8199 from Office of Economic Opportunity

Progress Report #5

November 1, 1969 - January 31, 1970

INTRODUCTION

Progress reported for the period from November 1, 1969 to January 31, 1970 consists of the following results:

1. Results from the analyses of the in-class monitoring procedure. These results indicate the significant differences among the four programs with respect to "Treatment Dimensions", which include both teaching techniques used by teachers and the amounts of similar activities engaged in by children. The results also include data on grouping in classes.

2. Results of analyses on the dependent variables, with summaries. (These data were briefly reported in the fourth progress report, but are included here for continuity.)

3. Preliminary report on results from regression analyses to explore the relationships between treatment dimensions and outcome variables.

4. Graphs on selected results from the video-tape monitoring procedure.

I. Treatment Dimensions

Method

Treatment (program) dimensions were assessed in two ways: (1) In-class monitoring with a time-sampling procedure which included both teacher and children, and (2) Video-tape monitoring of teachers only.

(1) In-class tallying was done by five monitors. The in-class tally procedure assessed seven categories of teaching techniques--manipulation of materials, verbal instruction, exemplification, motor activity, role-playing, physical guidance and conversation. The monitoring procedure was made as objective as possible so that very little interpretation or judgment was required for monitors to tally behavior in the various categories. Two estimates of these dimensions were used: (a) the amount of each technique relative to the total number of acts tallied. This is an index of the frequency of use of a given technique relative to other techniques used (designated "Cell/Row"). (b) the absolute amount of each technique as a proportion of the number of times tallying was done (number of 15-second periods). This is an index of how often a given technique was used regardless of the frequency of others (designated "Cell/Tally"). For example, a given teacher might use verbal instruction more than any other method when she does any teaching at all, and still do very little verbal instruction in terms of the time available to do so. On the other hand, a teacher might use a lot of verbal instruction, but also many other techniques, and the amount of verbal instruction relative to all of her methods would be low.

The in-class tally sheet also produced indices of the number of groups in classes, number of shifts in group size, relative proportion of kinds of groups--whether doing different things D/NF), the same thing (S), or engaged in a common enterprise (D/C), and total activity of all kinds tallied.

Analysis of variance was used to determine differences among programs on the in-class monitoring variables. Three different analyses were completed: a 3x4 analysis comparing the Bereiter-Engelmann, DARCEE, and Traditional programs which were in all four areas, a 2x4 analysis which included the Montessori program and compared all four programs in two areas, and a one-way analysis of variance which eliminated the area factor but also compared all four programs. Tukey's multiple comparison procedure was used to compare differences between the means. For the 2x4 analysis, only significant differences between Montessori and other programs are reported. The arcsine transformation was used to reduce positive skewness on the cell/tally and cell/row proportions. A Kruskal-Wallis one-way analysis of variance was used to detect program differences in number of changes in group size and type of group. The chi square test was used for the number of groups. For all differences reported as statistically significant, the probability level of the statistic used is .05 or less.

Results

(1) In-class Monitoring

(a) Teaching Techniques

Table 1 gives the absolute (cell/tally) percentage of four of the seven techniques tallied for teachers in the four programs. There were three techniques which were used so infrequently over all five of the two-hour monitoring periods that no analysis was possible. These were motor, physical guidance and role-playing. Motor and role-playing were included primarily in order to assess children, and their absence in teacher behavior was not particularly surprising, although Traditional teachers might be expected to use role-playing to some extent.

Verbal Instruction

Teachers

It was predicted that the Bereiter-Engelmann and DARCEE teachers would use verbal instruction significantly more often than teachers in the other two programs, and this was the case. Cell/tally percentages for both of these programs are significantly higher than in Traditional, and Bereiter-Engelmann is significantly higher than Montessori in the two areas where all four programs were located. Program order from most to least was Bereiter-Engelmann, DARCEE, Montessori, Traditional.

TABLE 1

MEANS FOR TEACHERS' BEHAVIOR IN FOUR PRESCHOOL PROGRAMS

Absolute Amount (Cell/Tally)[1]

Program	Manipulation	Instruction-Verbal	Exemplary	Motor Activity	Role-Playing	Conversation
Bereiter-Engelmann	.1706	1.1276[a]	0.9917[b]	Virtually zero	Virtually zero	.1436
DARCEE	.1220	1.0789[a]	0.5799			.2765
Montessori	.2971	0.8916	0.5563			.2503
Traditional	.2359	0.7895	0.5161			.2656

[1] Arcsine transformation

[a] Bereiter-Engelmann greater than Traditional and Montessori; DARCEE greater than Traditional.

[b] Bereiter-Engelmann greater than DARCEE, Montessori and Traditional.

Table 2 gives the relative percentage (cell/row) for the four techniques. In terms of the proportion of verbal instruction to all teaching methods tallied, DARCEE teachers used a significantly greater amount than Bereiter-Engelmann and Traditional.

Children

The percentage of the various categories tallied for children are shown in Table 3 (cell/tally) and Table 4 (cell/row). "Verbal instruction" for children may be thought of as recitation. The category collects all formal verbal behavior. Figure 1 shows that the absolute amount of verbal behavior is significantly greater for children in the Bereiter-Engelmann program than in the DARCEE and Traditional. In the two areas containing Montessori classes, Bereiter-Engelmann children also had significantly more verbal behavior than Montessori. The order of programs for children parallels the ordering for teachers; that is, Bereiter-Engelmann, DARCEE, Montessori, Traditional. This would be the case if the amount of recitation is a function of the amount verbally elicited by the teacher. Cell/row, or relative percentages, reflect the same order as absolute amounts. Relative to all acts tallied, Bereiter-Engelmann children and DARCEE children did significantly more recitation than Traditional. Bereiter-Engelmann was also significantly greater than DARCEE, and within two areas Bereiter-Engelmann was greater than Montessori.

Exemplification

Teachers

All teachers were expected to be high in exemplification; however, Bereiter-Engelmann is significantly higher than the other three programs in absolute amount and greater than DARCEE and Traditional in relative amount. This result explains the rather low cell/row percentage for verbal instruction in the Bereiter-Engelmann program. In this program verbal instruction is almost always accompanied by showing the children something (exemplification), usually a page in the teacher's manual. Thus, while the absolute amount of verbal instruction is high, it is not so high relative to other techniques. Exemplification was not expected to differentiate programs except that the kind of exemplification in Bereiter-Engelmann was expected to be almost entirely auditory and visual; relatively more sensorial stimulation in other modalities was expected in the other three programs. Checking this prediction requires content analysis which has not been done.

Children

For children, absolute percentages on exemplification do not

TABLE 2

MEANS FOR TEACHERS' BEHAVIOR IN FOUR PRESCHOOL PROGRAMS

Relative Amount (Cell/Row)[1]

Program	Manipulation	Instruction-Verbal	Exemplary	Motor Activity	Role-Playing	Conversation
Bereiter-Engelmann	.234	1.622	1.3907[c]	Virtually zero	Virtually zero	.1905
DARCEE	.201	1.952[b]	0.9314			.4352[d]
Montessori	.490	1.789	1.0067			.4352
Traditional	.501[a]	1.696	1.0057			.5356[d]

[1] Arcsine transformation

[a] Traditional greater than Bereiter-Engelmann and DARCEE.

[b] DARCEE greater than Bereiter-Engelmann and Traditional.

[c] Bereiter-Engelmann greater than DARCEE and Traditional.

[d] DARCEE greater than Bereiter-Engelmann; Traditional greater than Bereiter-Engelmann.

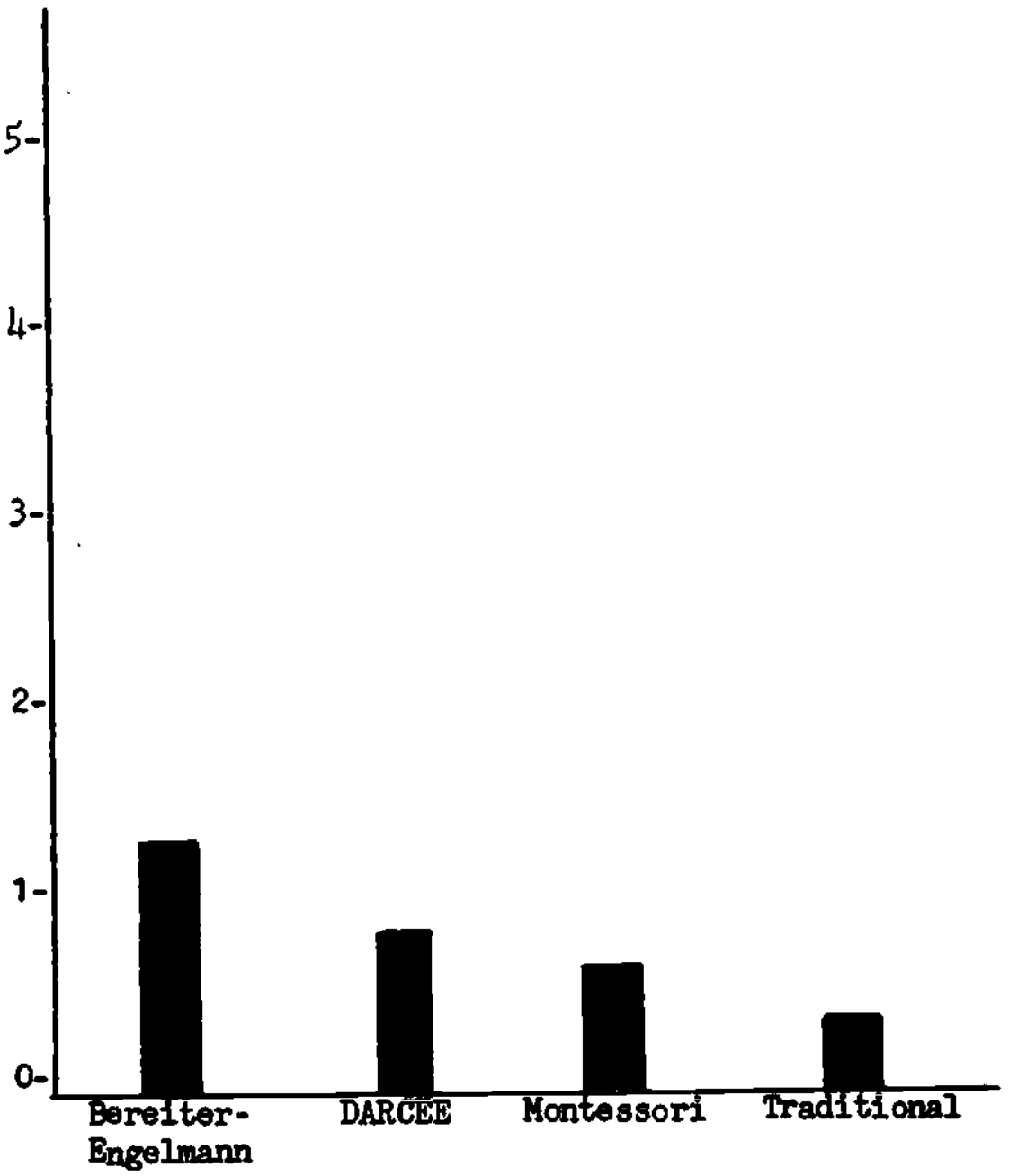

Figure 1. Children's verbal recitation per total number of acts tallied.[1]

[1] Arcsine transformation

TABLE 3

MEANS FOR CHILDREN'S BEHAVIOR IN FOUR PRESCHOOL PROGRAMS

Absolute Amount (Cell/Tally)[1]

Progr.	Manipulation	Instruction-Verbal	Motor Activity	Exemplary	Role Playing	Conversation
Bereiter-Engelmann	1.1589	0.9984[b]	0.3923	.3793	.1820	.3863
DARCEE	1.1597	0.5606	0.3492	.4271	.2403	.4827
Montessori	1.8142[a]	0.5030	0.1793	.2542	.2743	.6594
Traditional	1.4003	0.3007	0.3629	.3121	.6702[c]	.6600

[1] Arcsine transformation

[a] Montessori greater than Bereiter-Engelmann and DARCEE.
[b] Bereiter-Engelmann greater than DARCEE, Montessori, and Traditional.
[c] Traditional greater than Bereiter-Engelmann, DARCEE, and Montessori.

TABLE 4

MEANS FOR CHILDREN'S BEHAVIOR IN FOUR PRESCHOOL PROGRAMS

Relative Amount (Cell/Row)[1]

Program	Manipulation	Instruction-Verbal	Motor Activity	Exemplary	Role Playing	Conversation
Bereiter-Engelmann	1.3353	1.2117[b]	.5066	.4670	.2187	.4278
DARCEE	1.5622	0.6907[b]	.4676	.5346[c]	.2855	.6940
Montessori	1.9963[a]	0.5368	.1755	.2604	.2906	.6841
Traditional	1.5319	0.3306	.3957	.3317	.7271[d]	.6992

[1]Arcsine transformation

[a] Montessori greater than Bereiter-Engelmann.

[b] Bereiter-Engelmann greater than DARCEE, Montessori, and Traditional; DARCEE greater than Traditional.

[c] DARCEE greater than Montessori.

[d] Traditional greater than Bereiter-Engelmann, DARCEE, and Montessori.

differentiate programs, but DARCEE children are highest. In the two areas containing Montessori classes, DARCEE children had significantly higher cell/row percentages than children in Montessori and Bereiter-Engelmann classes. Tables 3 and 4 show that both relative and absolute percentages of exemplification were lowest in Montessori classes, though not significantly so.

Manipulation

Teachers

Bereiter-Engelmann teachers were expected to be low on manipulation, but they were not. Cell/tally percentages did not differentiate significantly but Montessori teachers were highest, as predicted. DARCEE teachers were lowest. Relative to other techniques, Traditional teachers did significantly more manipulation of materials as a teaching technique than teachers in the Bereiter-Engelmann and DARCEE programs (Table 2). This result was unexpected, but probably reflects the low incidence of other techniques in the Traditional program.

Children

Manipulation is more important as a technique in terms of its use by children than by teachers. Table 4 shows that Montessori children in the two areas containing these classes did significantly more manipulation of materials than did children in Bereiter-Engelmann and DARCEE classes. High incidence of manipulation by Montessori children was the expected outcome. Relative percentages are also higher for Montessori children than for Bereiter-Engelmann children--reflecting the fact that children in Montessori classes were usually manipulating materials, whereas those in the Bereiter-Engelmann were usually reciting. Children in the Traditional program were expected to be high in manipulation and they were second highest in absolute percentage--not significantly different from Montessori children in those two areas, as predicted.

Conversation

Teachers

Absolute amounts of teachers' conversation with children did not reach significance at the .05 level (Table 1) but the ordering is as predicted, from most to least, DARCEE, Traditional, Montessori, and Bereiter-Engelmann. These Cell/row percentages are similar and statistically significant, with DARCEE and Traditional being greater than Bereiter-Engelmann. Montessori teachers, however, had as much conversation with the children relative to other techniques as DARCEE teachers did.

Children

The category of children's conversation includes both conversation with teachers and with other children. The difference between greatest and least for cell/tally is not quite significant at the .05 level (Table 3), using the relatively conservative Tukey test, but Figure 2 shows that there was more conversation in the Traditional program than in Bereiter-Engelmann, as predicted. Surprisingly, there was almost as much in the Montessori classes as in the Traditional classes, and the amount in DARCEE classes was not especially high.

Role-playing

Children

It was predicted that role-playing by children should be highest in Traditional classes and least in Bereiter-Engelmann. This was the case. Traditional is significantly higher in both absolute (Table 3) and relative (Table 4) percentages than each of the other three programs. Figure 3 shows the absolute amounts of role-playing.

Motor

Children

We had expected motor behavior by children to differentiate the Traditional program from the others. However, this prediction was not borne out. There were no significant differences between programs in this category, though it is noteworthy that both relative and absolute percentages were highest in the Bereiter-Engelmann program.

(b) Grouping:

Number of groups. The size of any classroom limits the spatial separation of children and thus restricts the number of groups to the number of locations in which children can be physically isolated. Even though all 20 children were working alone, some would have to be at a table or in a corner of the room in physical proximity. The possible range of number of groups is therefore very narrow. The actual maximum for any class at any time was six. A chi square was calcualted on the frequency of occurrence of more than three groups versus three or less. The statistic was significant at the .001 level. Observed and expected frequencies are shown in Table 5, and are clearly in the expected direction.

Changes in group size. Changes in group size were computed as a ratio of changes to number of tally periods in order to eliminate the effects of number of groups on changes in size.

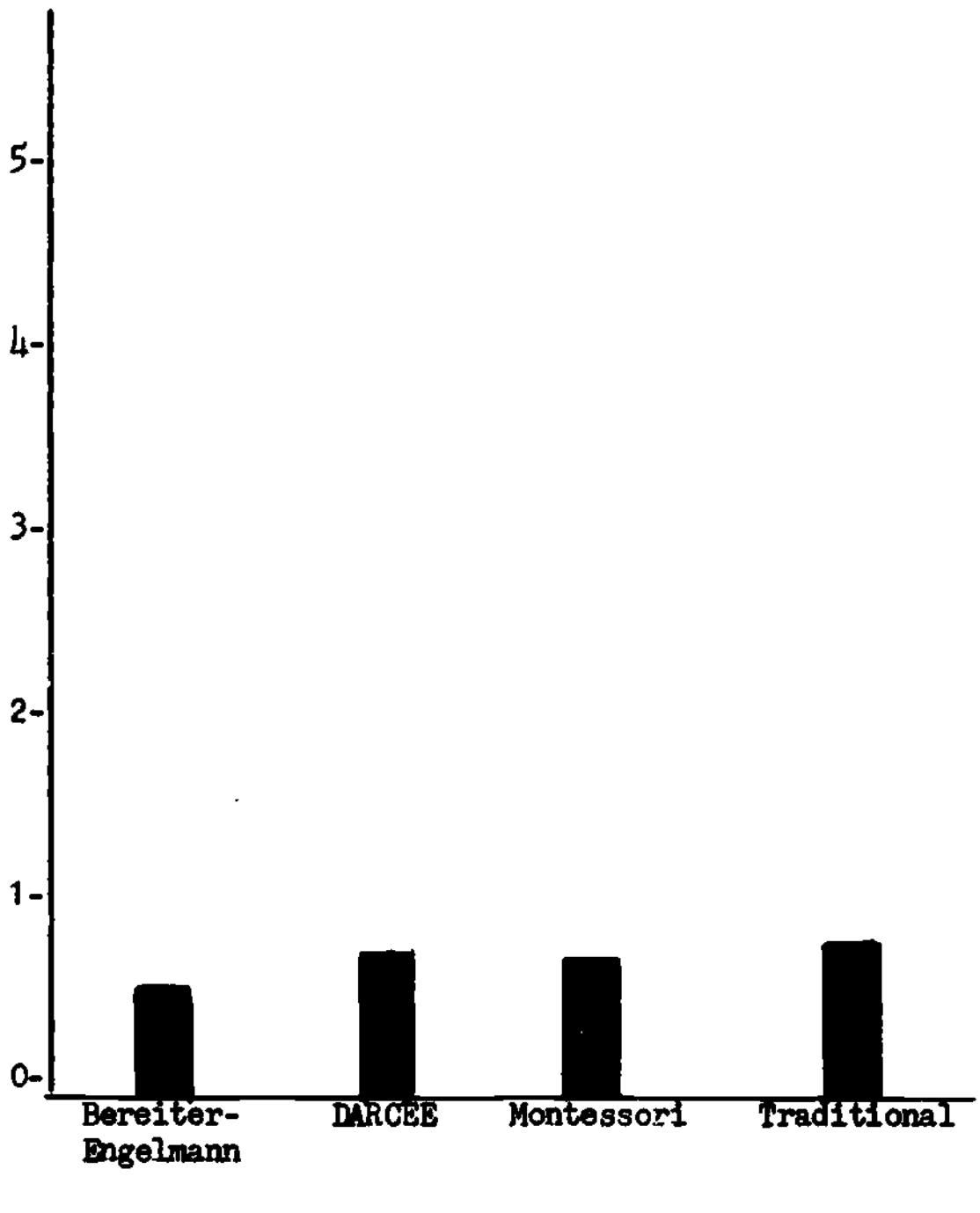

Figure 2. Children's conversation behavior per total number of acts tallied.[1]

[1] Arcsine transformation

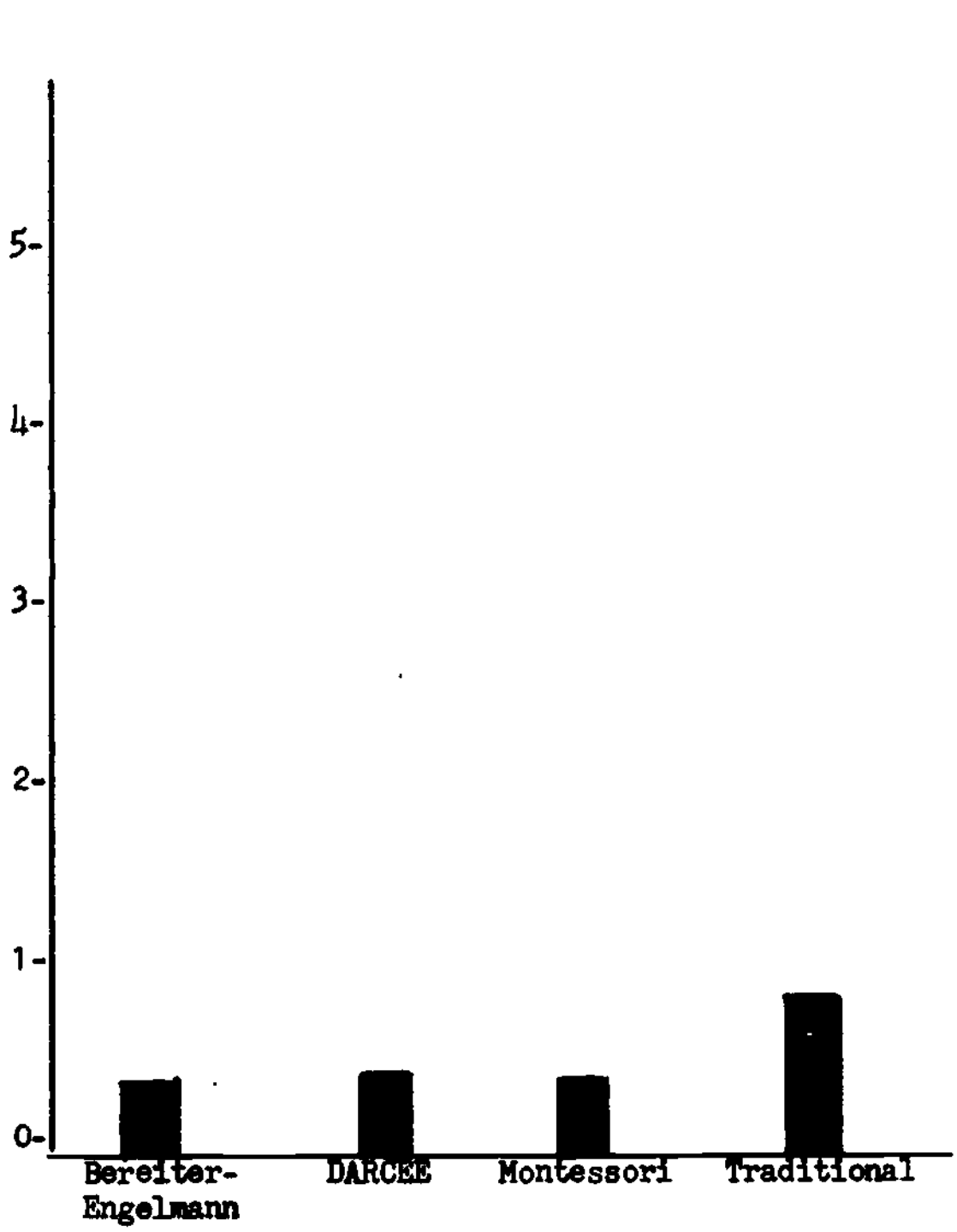

Figure 3. Children's role playing behavior per total number of acts tallied.[1]

[1]Arcsine transformation

TABLE 5

GROUPING IN FOUR PRESCHOOL PROGRAMS

	Programs			
	Bereiter-Engelmann	DARCEE	Montessori	Traditional
Number of Groups				
Mean frequency of more than three	8.00	4.00	16.00	11.25
Shifts in Group Sizes				
Mean shifts per tally period	.080	.079	.297	.286
Type of Group Activity				
	%	%	%	%
D/NF (Different/no group focus)	.058	.055	.177	.161
S (Same)	.877	.876	.779	.742
D/C (Different/common group focus)	.065	.068	.042	.096
Total Activity				
Mean ratio of acts per tally period	.5410	.4309	.3045	.2922

Table 5 shows that shifts in composition of groups were significantly less frequent in Bereiter-Engelmann and DARCEE than in Traditional. This reflects the predicted stability of groups in the more structured programs where children are instructed in three small groups. Group size changed as much in Montessori, however, as in Traditional.

Kinds of groups. Table 5 shows that groups of children who were all engaged in the same kind of activity (S) were significantly more frequent in Bereiter-Engelmann and DARCEE than in Traditional, as they should be if the programs were properly implemented. The converse of this is seen under D/NF where it is clear that groups consisting of children who were simply in physical proximity to each other but doing different things were significantly more frequent in Montessori and Traditional. Percentages under the D/C column reveal that there was very little cooperative effort toward a common goal or integrative play among these four-year-olds. We had predicted there would be more in Traditional classes, and though the absolute amount was slightly greater in these classes, the difference is not statistically significant.

Total Activity. It was predicted that teachers in Bereiter-Engelmann and DARCEE would be most actively engaged in teaching, and that Traditional and Montessori teachers would be less obtrusive in the classroom. Table 5 shows for each program the ratio of teaching techniques of any kind to the number of times tallying was done. As predicted, Bereiter-Engelmann and DARCEE teachers were doing more overt teaching than those in Traditional and Montessori.

In summary, a number of the program dimensions studied did differentiate the programs in predicted directions at a statistically significant level. Results not found were the predicted high frequency of motor (large muscle activity) by children in Traditional programs, and of exemplification (sensorial stimulation) by teachers in Montessori.

(2) Video-Tape Monitoring

Method

The procedure for monitoring teacher behavior from video-tapes was a modification and elaboration of the Bales Social Interaction Procedure. It assessed teacher behavior in three major areas: (1) what teachers were giving to children (information, stimulation, opinion, procedural information, etc.), and how this was done (verbally, modeling, etc.); (2) what teachers were eliciting from children; and (3) what feedback teachers were giving. A few other items, such as "out-of-contact", were also tallied.

Results

Analysis of these variables is not complete, but in many cases inspection reveals differences among programs of such magnitude that it seems safe to assume statistical significance. Knowledge-of-results (KOR) for correct responses, KOR for incorrect responses, contingent reinforcement(positive), requests for academic verbal performance, and requests for imitation are variables which appear most likely to differentiate programs. Figures 4, 5, 6, 7, and 8 show these mean frequencies.

II. Treatment Effects

Method

Treatment (program) effects were assessed in five areas: cognitive, social, motivational, perceptual, and achievement. Analysis of the dependent variables followed the same procedure as was used with the in-class monitoring variables: a 3x4 analysis, a 2x4 analysis, and a one-way analysis which also included the control group. Since analysis of covariance on the spring test, using the fall test as a covariant, was inappropriate for many variables, separate analyses of variance were computed for the fall and spring tests. Tukey's multiple comparison procedure was again used to determine differences between means.

However, the first analyses left some questions still unanswered; primarily those of sex differences and differences in the amount of change from fall to the spring test. Thus a repeated measures analysis which also included the sex factor was used. This analysis was chosen because it provides interactions not directly provided by analysis of change scores and has greater precision. The repeated measures analysis has just been completed and the results seem to support previous conclusions. In most cases there were pre-post differences indicating change from the fall to the spring tests for all subjects. Several pre-post by program interactions occurred, indicating differences in the amount of change by program. These differences support previous findings.

Results

(1) Cognitive Measures

Three tests were given to all subjects, including controls, in the fall (after 6-10 weeks of school) and again in the spring (after 6 months of school). These were the Stanford-Binet, the Preschool Inventory, and the Quick Picture Vocabulary Test. In addition the Peabody Picture Vocabulary Test was given in the spring to a sample of six children from each experimental class. Separate analyses of variance were made of fall and spring tests. No differences mong programs were found at either testing on the Quick Test. On the Binet, the DARCEE program was

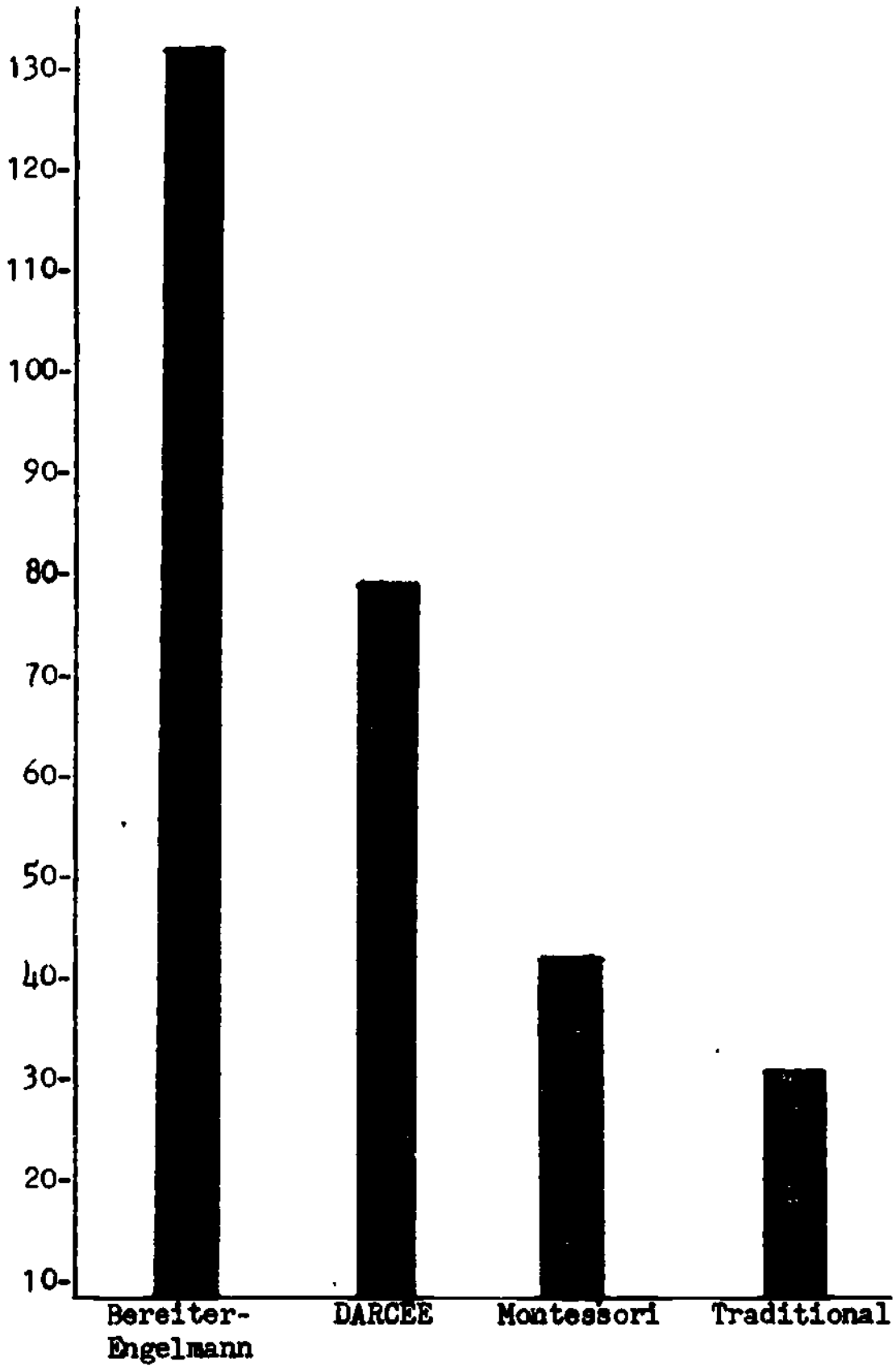

Figure 4. Mean number of times teachers provided children with knowledge of results. (Correct responses.)

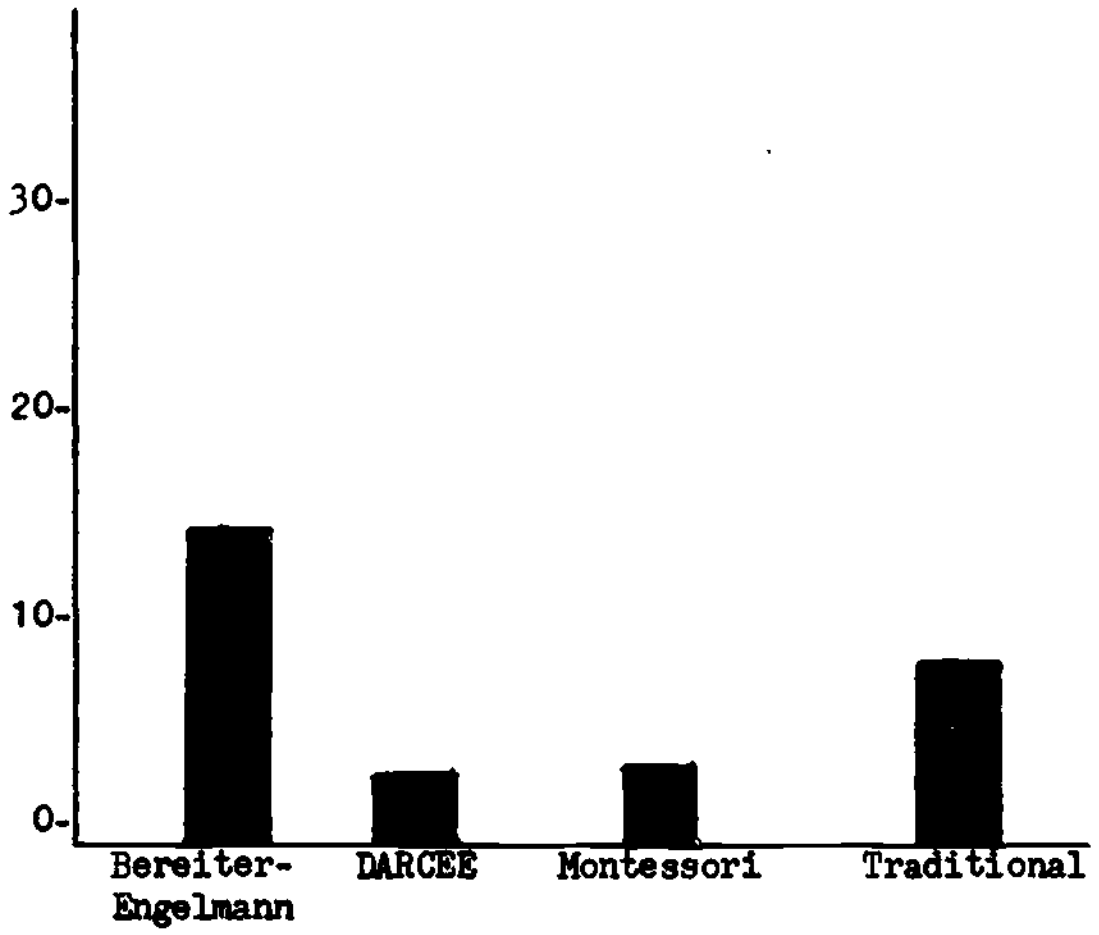

Figure 5. Mean number of times teachers provided children with knowledge of results. (Incorrect responses.)

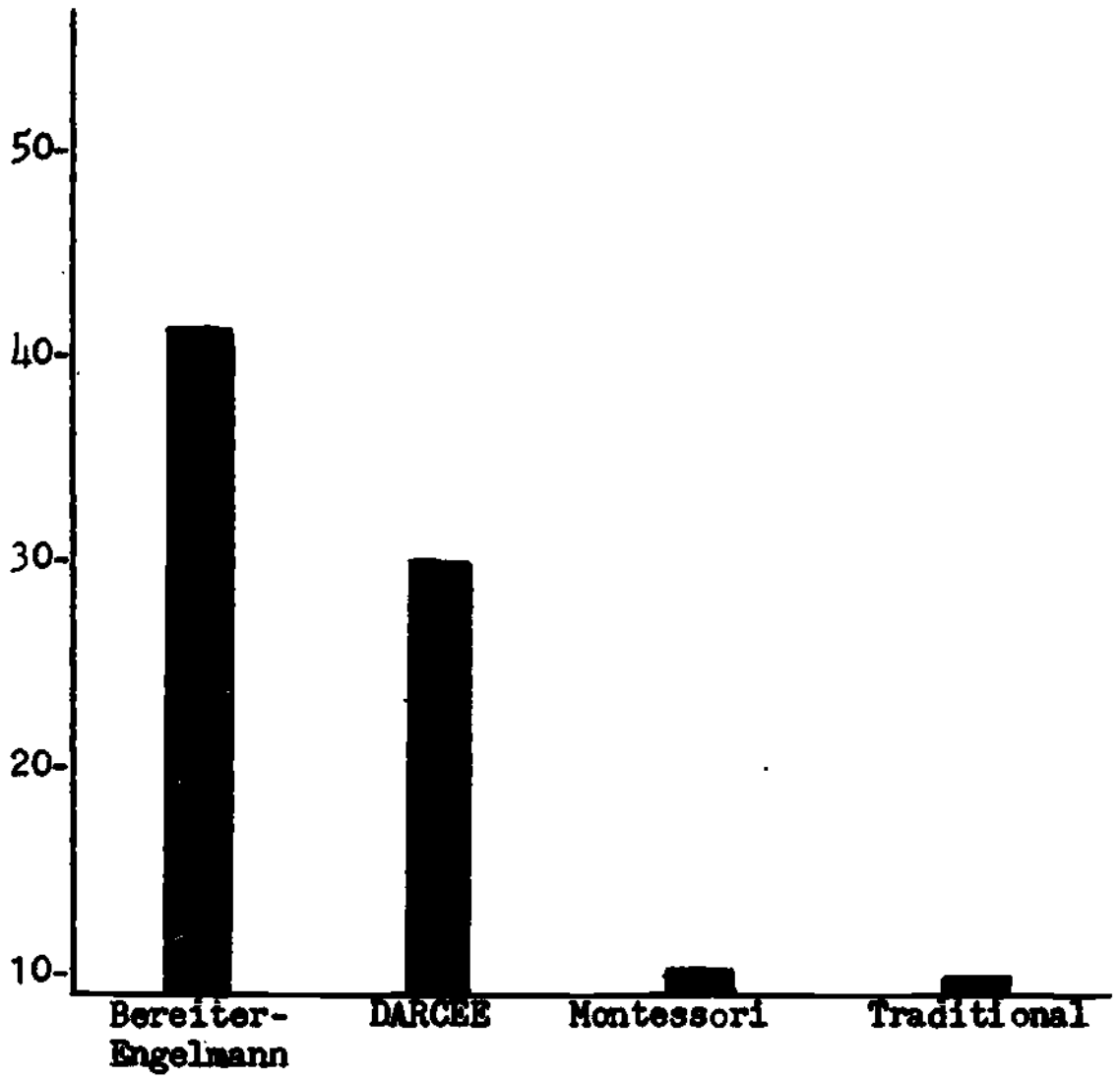

Figure 6. Mean number of times teachers provided reinforcement for correct or appropriate behavior.

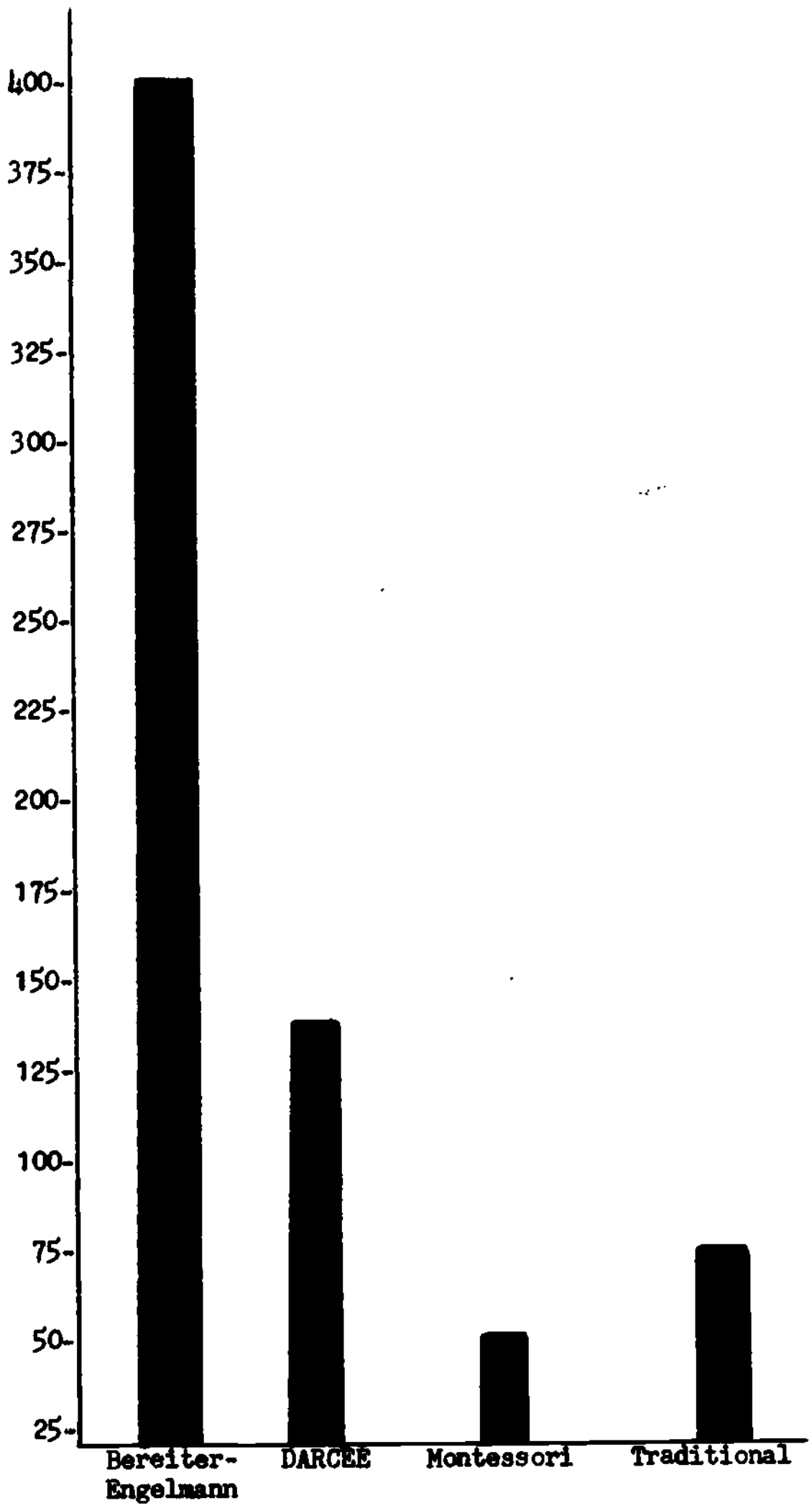

Figure 7. Mean number of requests by teachers for children's academic verbal performance.

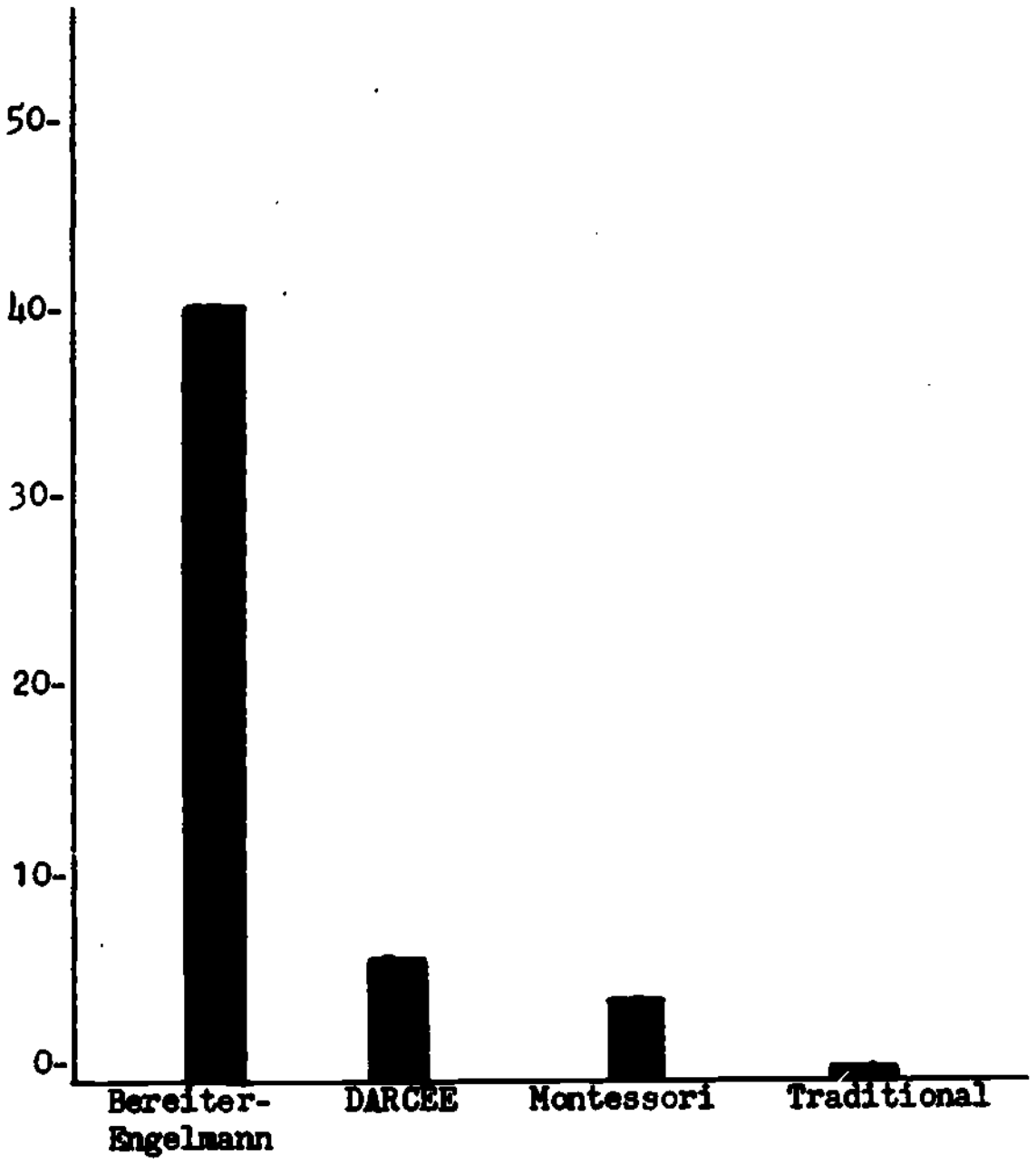

Figure 8. Mean number of requests by teachers for children to imitate teachers' performance.

significantly higher than the Traditional program on the fall test; on the spring test, both DARCEE and Bereiter-Engelmann were higher than controls (Table 6). On the Preschool Inventory, DARCEE was above Traditional in the fall and in the spring; in the spring both DARCEE and Bereiter-Engelmann were higher than controls (Table 7). Spring testing with the Peabody Picture Vocabulary Test did not reveal any significant program differences.

In summary, the DARCEE and Bereiter-Engelmann programs on the basis of preliminary analysis appear to have significantly affected cognitive functioning as measured by the Stanford-Binet and the Preschool Inventory. Traditional and Montessori were not significantly different from controls.

(2) Social-Emotional Measures

Two rating scales were used in the fall and again in the spring to assess effects of programs on social and emotional behavior--the Behavior Inventory (completed by both teachers and aides) and the Face Sheet of the Binet (completed by both Binet testers and Preschool Inventory testers). Analyses of the Face Sheet are not complete, and analyses of the Behavior Inventory ratings are complete only for teachers' ratings. These results must therefore be interpreted with caution, as they are based entirely on teachers' ratings. Results of Preschool Inventory testers' ratings on the Binet Face Sheet, however, appear to be consistent with Behavior Inventory results on achievement motivation. Measures for all factors on the Behavior Inventory are shown in Table 8.

Only one significant program effect appeared on the fall ratings--on the Independence factor, where Traditional children were rated significantly more independent than were DARCEE children.

In the spring, a number of program differences appeared. DARCEE children were rated significantly higher in Verbal-Social-Participation than were children in Bereiter-Engelmann and Montessori; children in Traditional were rated higher than those in Montessori.

On Timidity, DARCEE children were rated significantly better than those in Bereiter-Engelmann and Traditional.

Independence showed no program differences in the spring. In Aggression, Bereiter-Engelmann children were rated significantly better than Traditional children.

Achievement factor results are presented under motivational measures.

In summary, according to teachers' ratings after 6 months, children in the DARCEE program were significantly better in Verbal-Social-Participation and less timid than was the case in

TABLE 6

STANFORD-BINET TEST - FORM L-M

MEANS BY PROGRAMS

Program	Fall	Spring
Bereiter-Engelmann	93.25	99.52[b]
DARCEE	96.02[a]	97.56[b]
Montessori	91.50	96.34
Traditional	89.35	95.02
Controls	89.21	90.00

[a]DARCEE greater than Traditional. [b]Bereiter-Engelmann and DARCEE greater controls.

TABLE 7

PRESCHOOL INVENTORY TEST

MEANS BY PROGRAMS

Program	Fall	Spring
Bereiter-Engelmann	26.33	39.06[b]
DARCEE	28.92[a]	40.98[b]
Montessori	25.21	37.55
Traditional	24.36	35.98
Controls	28.29	33.18

[a]DARCEE greater than Traditional. [b]DARCEE and Bereiter-Engelmann greater than controls; DARCEE greater than Traditional.

TABLE 8

BEHAVIOR INVENTORY

MEANS BY PROGRAMS

Factor	Programs							
	Bereiter-Engelmann		DARCEE		Montessori		Traditional	
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
Achievement	11.95	11.85	11.92	13.29[a]	11.28	11.28	12.65	12.38
Verbal-Social	10.03	12.07	11.51	14.39[b]	10.41	10.94	11.83	12.92[b]
Timidity	11.97	12.10	11.61	13.97[c]	12.97	13.06	12.58	12.54
Independence	11.73	11.95	11.05	13.19	12.38	12.72	12.75[d]	13.10
Aggression	13.34	13.78[e]	12.76	13.20	12.50	12.66	13.40	12.21

[a]DARCEE is greater than Montessori.

[b]DARCEE is greater than Bereiter-Engelmann and Montessori; Traditional is greater than Montessori.

[c]DARCEE is greater than Bereiter-Engelmann and Traditional

[d]Traditional is greater than DARCEE.

[e]Bereiter-Engelmann is greater than Traditional.

the other programs; Bereiter-Engelmann children were significantly less aggressive.

(3) Motivational Measures

Assessments of motivational variables were made by tests in four areas: curiosity (Curiosity Box), persistence and resistance to distraction (Replacement Puzzle), inventiveness or initiative (Dog-and-Bone), and achievement motivation (Behavior Inventory and the Binet Face Sheet).

On the spring test, the DARCEE, Bereiter-Engelmann, and Montessori children were significantly higher than controls in curiosity activity (exploration of the box, Table 9). The DARCEE children were more task persistent and more resistant to distraction than controls (Table 10). In inventiveness, DARCEE children scored significantly higher than those in Bereiter-Engelmann and Traditional programs (Table 11). On achievement motivation, as rated by teachers on the Behavior Inventory, DARCEE children were higher than those in Montessori (Table 8). Face Sheet analysis has not been completed.

In summary, analyses to date suggest that the DARCEE program had considerable impact on children's motivation to achieve, persistence, resistance to distraction, initiative, and curiosity. Bereiter-Engelmann and Montessori also were significantly high in curiosity.

(4) Perceptual Measures

Two instruments were used to assess perceptual functioning: the Embedded Figures Test and the Wepman Auditory Discrimination Test. On Embedded Figures there was moderate improvement for all groups from fall to spring testing but no differences among groups, even between experimentals and controls. One area-by-program interaction occurred but this would seem to be entirely due to the failure of one class to improve on the test.

The Wepman Auditory Discrimination Test proved to be an inappropriate test as standardized for these four-year-olds. A large number of tests were invalid because of response sets and testers reported their impressions that children simply were not able to understand the instructions. The use of the terms "same" and "different" is probably confusing as well as the testing procedure, which calls for the child to be placed with his back to the examiner. The results cast doubt on the validity of even those tests on which no response set was found. Therefore no analyses were made except that the percentage of valid tests for each group on the spring testing was calculated and these percentages are as follows: Bereiter-Engelmann, .83; Montessori, .76; Traditional, .66; DARCEE, .63, controls, .50. The large difference between the percentage of valid tests for children in the Bereiter-Engelmann program and the controls is suggestive,

TABLE 9

CURIOSITY BOX TEST - ACTIVITY

MEANS BY PROGRAMS

Program	Fall	Spring
Bereiter-Engelmann	17.72	18.06[b]
DARCEE	14.98	17.81[b]
Montessori	19.76	18.67[b]
Traditional	17.32	17.15
Controls	16.59	14.09

[b]Montessori, Bereiter-Engelmann, and DARCEE are greater than controls.

TABLE 10

REPLACEMENT PUZZLE TEST

MEANS BY PROGRAMS

A

Persistence

Program	Fall	Spring
Bereiter-Engelmann	21.21[a]	22.13
DARCEE	20.41[a]	22.75[b]
Montessori	19.72	22.07
Traditional	20.79[a]	22.35
Controls	17.47	20.65

[a]Bereiter-Engelmann, DARCEE, and Traditional are greater than controls.
[b]DARCEE is greater than controls.

B

Resistance to Distractor

Program	Fall	Spring
Bereiter-Engelmann	9.73	9.48
DARCEE	8.35	10.10[b]
Montessori	9.62	7.72
Traditional	8.67	8.96
Controls	8.18	7.06

[b]DARCEE is greater than controls.

TABLE 11

DOG AND BONE (INVENTIVENESS)

MEANS BY PROGRAMS

Program	Fall	Spring
Bereiter-Engelmann	3.21	4.19
DARCEE	3.58	6.36[b]
Montessori	4.06	5.61
Traditional	3.30	4.23
Controls	4.06	4.97

[b]DARCEE is greater than Bereiter-Engelmann and Traditional.

but interpretation should probably await further testing.

In summary, no program differences were found for perceptual functioning, a result which seems in the case of auditory discrimination to be due to the nature of the test used.

(5) Achievement Measures

Four tests were given to a sample of children from each class at the end of the school year in order to assess achievement in specific areas. These were: Arithmetic, Parallel Sentence Production, Expressive Vocabulary, and Basic Concept Inventory. Two of these - Arithmetic and Parallel Sentence Production - revealed significant program effects.

On Arithmetic, the children in the Bereiter-Engelmann program scored significantly higher than those in Traditional and Montessori. DARCEE children were significantly better than Traditional (Table 12).

Parallel Sentence Production significantly differentiated Bereiter-Engelmann and Traditional children in favor of Bereiter-Engelmann. Table 13 shows that the mean for DARCEE children was 13 points higher than the mean for Traditional children, although this difference is not quite significant at the .05 level.

The Basic Concept Inventory was developed by Engelmann in connection with the Bereiter-Engelmann program. Variability within programs was quite high. Numerous interactions would be expected if classes differed and if the test is accurately identifying educational deficiencies. The three parts of the test which assess Basic Concepts, Statement Repetition and Comprehension, and Pattern Awareness have not been examined separately. Total score did not differentiate programs.

In summary, as predicted, children in Bereiter-Engelmann classes who were trained in arithmetic scored significantly higher on a test of it. DARCEE children, who also had some numerical training, scored second highest. These two programs also stressed production of complete sentences and appropriate syntactical form. The effects of this training are shown in results in Parallel Sentence Production. Thus, programs designed to teach special skills did succeed in doing so to a significant degree.

III. Relations Between Treatment Dimensions and Treatment Effects

As a preliminary analysis, multiple regressions were done to assess the relationships between teaching techniques obtained from the in-class monitoring procedure and change scores (fall-spring) on each of the dependent variables. Scores on each of the dependent variables were available for each subject; however, scores on teaching

TABLE 12

ARITHMETIC TEST

MEANS BY PROGRAMS

Program	Fall	Spring
Bereiter-Engelmann	----	17.75[b]
DARCEE	----	13.42[b]
Montessori	----	8.47
Traditional	----	6.67

[b]Bereiter-Engelmann is greater than Traditional and Montessori; DARCEE is greater than Traditional.

TABLE 13

PARALLEL SENTENCE PRODUCTION TEST

MEANS BY PROGRAMS

Program	Fall	Spring
Bereiter-Engelmann	----	95.88[b]
DARCEE	----	90.58
Montessori	----	84.83
Traditional	----	77.88

[b]Bereiter-Engelmann is greater than Traditional.

techniques were available only for classes. Several points should be made regarding this analysis.

1. The program variable was not included in the regression analysis since the purpose was to examine the relationship between the activities assessed by in-class monitoring and changes in the dependent variables independent of knowledge of the program itself. It was already established that programs differed on the in-class monitoring variables and also on the dependent variables. Thus, the effect of including program in the regression analysis would be to attenuate the beta weights for those classroom variables highly correlated with programs.

2. Only those teaching techniques which were included on the in-class tally sheet were represented in this analysis. In view of the significant differences among programs, and also to some extent among teachers within programs, on the video-tape variables, (such as knowledge of results, reinforcement, and requests for imitation), it is expected that the inclusion of these variables in future regression analyses will augment the multiple correlations.

3. A number of other variables which would probably be related to change scores were also not included. The most likely of these would be the pre-test score on the dependent variables. Other factors outside the purview of this experiment such as home life, contribution of parents in extending the goals of the preschool, children's health, the interaction between the child's personality and programs might also have an effect on the amount of change.

4. Since the values of the predictor variables were means for classes and the values of the dependent variables were individual scores, high multiple correlation coefficients were not expected. If the variability of change scores within each of the classes was small, then the multiple correlation could be expected to be high. However, if the variability of the change scores within the classes was great, (which might be the case, for example, if the teacher paid more attention to some children than others), then the multiple correlations between the in-class variables and the dependent variables would be an underestimation of the actual relationship.

Multiple Rs for the predictor variables ranged from .23 to .42. Partial correlations between the criterion variable and the given predictor, holding the other predictors constant, ranged from -.29 to .31. Detailed analysis and interpretation of these results have not been completed.

Analyses in Process

Some additional results from data obtained from the first year of the study have not been reported but analyses are in process and results on many of them will be included in the Annual Report for 1969-70. They consist of the following:

1. Sex Differences

A repeated measures analysis of variance on dependent variables has been completed and comparisons between programs is in process. This analysis includes the sex variable. There were no sex main effects, indicating that sex was not an overriding factor; however, there were some interactions between sex and other variables.

2. The Binet Face Sheet

Results from the assessment of children by testers using the Face Sheet of the Stanford-Binet have not been reported in detail but these results are now available. In addition, the Face Sheet of the Binet was completed by both Binet testers and Preschool Inventory testers. Comparison between ratings for these two groups of testers will be made and correlations will be reported between Face Sheet ratings and test scores. This portion of the research replicates a study by Hess, et. al. (1966). These investigators found a substantial correlation between Face Sheet ratings and Binet IQ which may or may not be due to Binet testers' ability to estimate children's IQs and thus produce bias in their ratings.

3. DARCEE Sub-Study

Approximately one-half of the mothers of the children in the DARCEE program were visited once a week by the assistant teachers (home visitors) who took them materials and instructed the mothers regarding their use. Comparisons between visited and not-visited children have not been done with the exception of scores on the Stanford-Binet where spring test scores do not differ. Change scores as a function of this variable have not been compared on the Binet or any of the other dependent variables. Other variables may reveal the effect of this work with the mothers.

4. Controls: Waiting list vs. others

Comparisons are being made between controls obtained from the waiting list for Head Start and those recruited otherwise. It is often suggested that children whose parents voluntarily take advantage of special educational programs differ from the "hard-core" poor who are more difficult to reach. It can already be stated that in this study controls who registered for Head Start but could not participate because classes were full do not differ in Binet IQ from controls located by inquiries in the community. They may, however, differ on other variables.

5. Teacher-Aide Differences on Behavior Inventory

Both teachers and aides in each class completed the Behavior Inventory on the children and it will be of interest to determine whether there are significant differences between their ratings. Teachers had more training in their respective curricula than the aides; thus the variance of ratings within programs for the two raters may throw some light on the question of the extent to which the training programs changed the teacher's perceptions of children in respect to the factors assessed by this instrument. Amount of change seen by the two kinds of raters will also be of interest.

6. Regression Analysis of In-Class Tally Variables

Detailed results of the multiple regressions between variables assessed by the in-class monitoring procedure and the dependent variables will include partial correlations between absolute amounts of teaching techniques and dependent variables, absolute amounts of techniques used by children and the dependent variables, as well as correlations between the relative amounts of these variables and the dependent variables. Results obtained so far indicate that it may be possible to extract meaningful patterns of relationships between the activities of the teachers and children and the changes which occurred on the tests used.

7. Demographic Data

Two forms were used to obtain information regarding family size, mobility, educational levels of parents, occupations, facilities in the home, and a large number of other factors. Comparisons between programs and between areas will be of interest in respect to interpretation of results from both the pre-kindergarten and the kindergarten years and confirmation that random assignment of children was successful. In addition some of these data may be entered as predictor variables in future regression analyses.

8. Comparison of Wepman and CADI Tests of Auditory Discrimination

The failure of the Wepman Auditory Discrimination Test to provide enough analyzable data for an assessment of auditory discrimination in our experimental group resulted in the design of an additional study to compare the Wepman with the California Auditory Discrimination Index[1]. Both tests were given to the middle-class control group (N=48) and a group of four-year-old children enrolled in Head Start for the 1969-70 school year (N=48). The Head Start children were randomly selected from classes in the same schools from which the experimental sample was drawn the previous year. Sequence of administration and testers (two) were counter-balanced. This study

[1]UCLA Preschool Research Projects, Dr. Carolyn Stern, Director
1019 Gayley Ave., Los Angeles, California 90024

has been completed and analysis of variance will be used to analyze results. This will provide additional information about the CADI which has not been widely used and will further clarify the question of whether difficulties with the Wepman were due to deficit in auditory discrimination among disadvantaged children or the fact that the test is simply inappropriate for this age level. Instructions and procedure on the Wepman were modified somewhat in order to minimize problems of administration.

DISCUSSION

1. The fact that programs did differ from each other in many of the expected dimensions provides at least a partial answer to Jensen's question regarding the exportability of programs. There was no one on the research staff who was in any way involved with the development of any of the programs or who had any bias toward or against any particular program. Consultants visited with their teachers only twice during the entire year. The amount of supervision provided for these teachers was considerably less than that provided in programs such as Follow-Through. The only program for which relatively constant supervision was provided was Montessori, in which a local supervisor spent one day per week with each teacher.

A consultant's evaluation form filled out on each teacher reveals that the consultants rated programs in every case slightly above the mid-point in consideration of the limitations involved (Table 14).

Results from in-class monitoring are more convincing, since they present a random sample of what the teachers and children were doing in the classrooms throughout the year. These results indicate that despite considerable variation among teachers within each program, most of the salient characteristics of the programs which it was feasible to assess in this manner were clearly present, and to a sufficient extent to produce greater homogeneity within programs than between programs.

On the basis of the in-class monitoring, it can be concluded that Bereiter-Engelmann and DARCEE teachers did more teaching of all kinds, and used more verbal instructions than Montessori and Traditional teachers; exemplification, as a technique, was significantly high in Bereiter-Engelmann; teachers in Traditional classes used manipulation of materials more than any other technique. In Bereiter-Engelmann and DARCEE classes children worked in groups, engaged in the same kind of activity; in the Montessori and Traditional classes children were engaged in different activities, either individually or moving around in various groups, the composition of which fluctuated rapidly. Children in Bereiter-Engelmann and DARCEE classes did more reciting, children in Montessori more manipulating of materials, and children in Traditional more role-playing. Conversation among children was highest in Traditional classes, though this analysis is not quite statistically significant at the .05 level.

It is clear, however, that these programs were not prototypes of the originals in every respect. Bereiter-Engelmann and DARCEE programs were designed for five-year-olds and some modification was necessary simply because they were used with four-year-olds. Montessori classes should contain three-, four-, and five-year-olds rather than a homogeneous group of four-year-olds. In addition, we did not find greater motor activity by children in Traditional, nor greater sensorial stimulation by Montessori teachers. Motor behavior was quite variable within programs, which might account for the lack of significant program differences, but our impression is that the Traditional children were not actually engaging in any more of such activities than was the case in other programs. With regard to sensorial stimulation in Montessori classes, it may be that more of this

TABLE 14

CONSULTANTS' RATING OF PROGRAMS[1]

("Not at all" (0) to "Best possible" (10))

	PROGRAMS (Mean for all classes)			
	Bereiter-Engelmann	DARCEE	Montessori	Traditional
Program Aspect Evaluated				
Teaching Techniques	5.87	7.28	6.25	9.17
Materials	7.50	7.00	7.50	10.00
Principles	5.62	7.86	6.50	9.00
Context	5.50	7.71	7.00	8.50
Selection of Activities	5.00	7.86	7.25	8.83
Content	8.50	6.71	6.50	9.17
Facilities	7.87	8.29	5.00	3.67
Progress of Children	5.86	7.25	6.25	-
Classroom events typical	7.00	7.43	6.50	8.67
Extent to which a demonstration	5.37	7.71	6.25	10.00
Mean - All Categories	6.33	7.51	6.50	8.56

[1] Considering limitations (Relative Criterion)

took place during the first three months of the school year when we were not monitoring. In any case, Montessori teachers, like those in all programs, used verbal instruction more than they did exemplification. When kinds of exemplification are examined, further explanations may be possible.

The result of significant differences among program dimensions also answers the question whether different teaching techniques and actual classroom behavior resulted from these different preschool training programs - they did. Teachers in all programs except Montessori had previously taught in the Traditional program style. Yet their classroom activities after training resembled those of their counterparts in the same program more than they resembled those of teachers in other programs.

The fact that teachers' behaviors can be altered in significant ways as a result of 4 to 8 weeks of training seems to have important implications for the preparation of preschool teachers. Our teachers had a variety of backgrounds and levels of education, and certainly differed greatly in personality and style. Sigel (1969) has suggested that the most important factor in the success of any educational program may be, not the teachers' personalities per se, but the degree of understanding of, involvement with, and commitment to, the programs used. He also suggests, "Proper and appropriate teacher training and programming, both pre-service and in-service, can create the kinds of attitudes and orientations that are necessary for the diffusion of educational innovations." (Seminar #6, p. 6). Our results so far support the notion that regardless of differences in personality and formal education the preschool teacher can be greatly influenced in her teaching methods by training which involves specific procedures, and provides the teacher with unambiguous instructions regarding techniques. Some time ago, Thompson (1944),using teachers with extensive backgrounds in preschool education and above average children, also found that training programs had significant effects on teacher behavior. His results with a different population are thus supported and extended. Information on teacher personality and other data on teachers which is available but not yet analyzed may provide clarification.

We are unable to assess the influence of curricular materials on teachers' behavior, but it is our impression that materials do have coercive effects, particularly when the teacher's role is programmed in minute detail, as is the case in Bereiter-Engelmann.

2. The number of significant treatment effects on outcome variables is gratifying, particularly since two of the more structured programs--Bereiter-Engelmann and DARCEE--produced effects which are consonant with their goals. Results from the Montessori and Traditional classes are disappointing, especially since children in these classes did not compare favorably on those variables selected to assess the goals of these programs. With respect to the Hawthorne effect, it should be noted that the Traditional or regular Head Start program , which is sometimes evaluated as a control group, was included as an experimental group in this study. Thus, superiority of other programs, where found, cannot be attributed to the greater enthusiasm or dedication of teachers which might result from being a part of an experimental group.

Analyses are not complete, however, even on data already available and it would therefore be premature to draw firm conclusions. Both the Montessori and Traditional programs are oriented toward long-term rather than short-term goals and in neither case is there an attempt to move all children in the same direction so far as their development is concerned. Therefore, evaluation of their success cannot be made with confidence until results from later testing become available. And finally, the stability of all program effects over time is perhaps of greater importance than their immediate impact.

3. Results from the regression analyses completed so far appear promising with respect to providing hypotheses regarding the relationships between effective components of these programs and the dependent variables. Detailed examination of the results from the multiple regressions will provide partial correlations between teaching techniques used by teachers and children, and with respect to absolute and relative amounts of these techniques. Results obtained so far indicate that it may be possible to extract meaningful patterns from the data.

REFERENCES

Hess, R.D., Kramer, D., Slaughter, D., Torney, J., Berry, D., and Hull, E. "Techniques for Assessing Cognitive and Social Abilities of Children and Parents in Project Head Start", Report on Research Contract OEO-519 with Office of Economic Opportunity, University of Chicago, July, 1966.

Jensen, A. R., "How Much Can We Boost I.Q. and Scholastic Achievement?" Harvard Educational Review, 1969, 39, pp.1-123.

Sigel, I.E. "Analysis of Teacher Role in Educational Intervention Programs", in Critical Issues in Research Related to Disadvantaged Children, Edith Grotberg, Ed., Princeton: Educational Testing Service, 1969.

Thompson, G.G. "The Social and Emotional Development of Preschool Children Under Two Types of Education Programs", Psychological Monographs, 1944, 56, No. 5, (Whole No. 258).